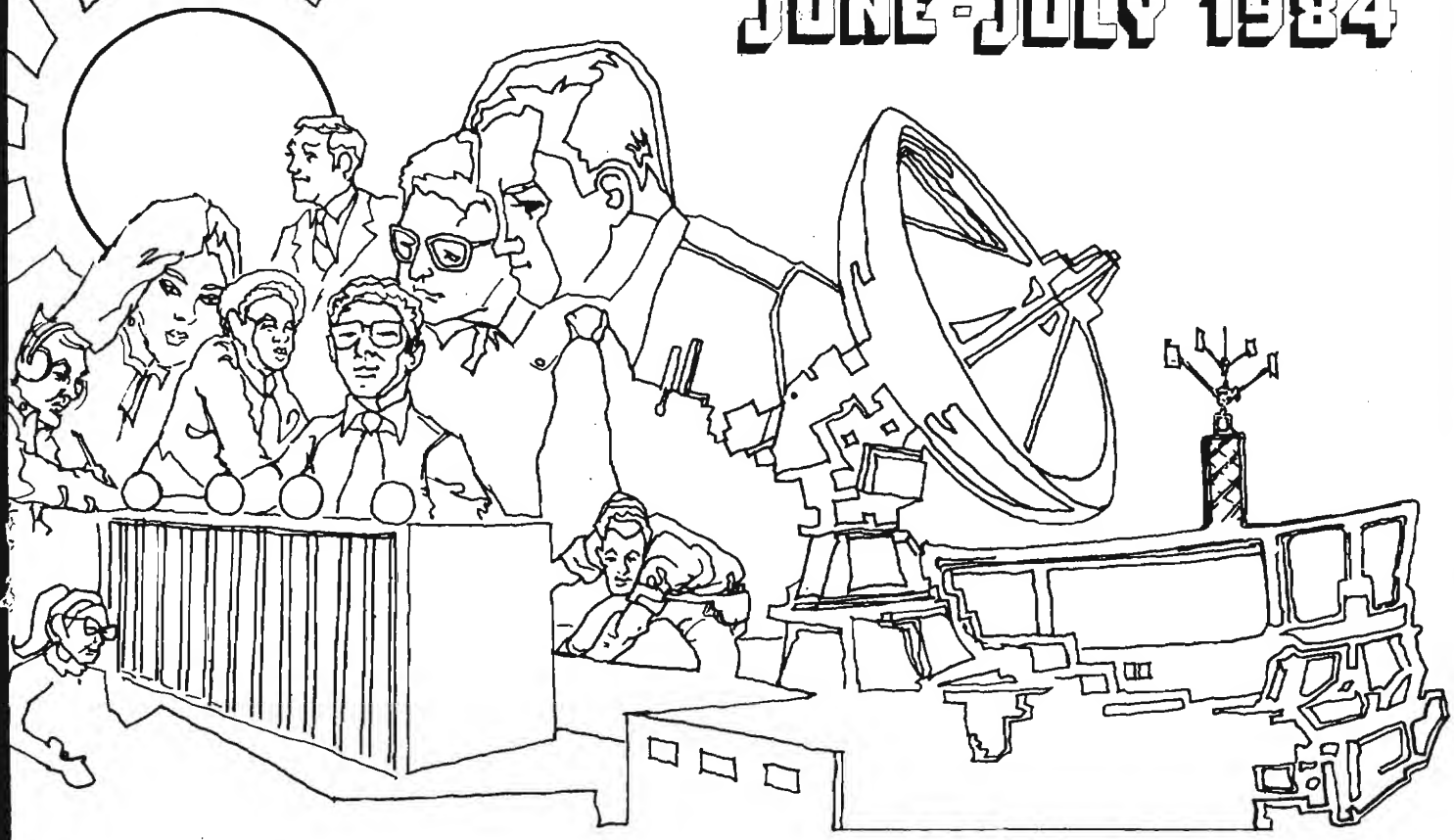
~~TOP SECRET~~

NATIONAL SECURITY AGENCY
FORT GEORGE G. MEADE, MARYLAND

CRYPTOLOG

JUNE-JULY 1984



TRAFFIC ANALYSIS: A CURRENT PERSPECTIVE (U).... [redacted] 1
ESPIONAGE AS A TOOL
OF TECHNOLOGY TRANSFER (U)..... [redacted] 8
INTELLIGENCE ANALYSIS
IN THE COMPUTER AGE (U).....Jack Gurin.....12
NAMING SOVIET CITIES (U).....16
THE BOOKBREAKER (U).....Stuart Buck.....17
WORD PROCESSING PLAIN AND SIMPLE (U)..... [redacted] 18
UNLESS TEXTS HANG TOGETHER, LINGUISTS
WILL ALL HANG SEPARATELY (U)..... [redacted] 25
NSA-CROSTIC NO. 56 (U)..... [redacted] 28

P.L. 86-36

~~THIS DOCUMENT CONTAINS CODEWORD MATERIAL~~

~~TOP SECRET~~

~~CLASSIFIED BY NSA/CSSM 123-2~~
~~DECLASSIFY ON: Originating~~
~~ration Required~~

CRYPTOLOG

Published by P1, Techniques and Standards

VOL. XI, No. 6-7 JUNE-JULY 1984

For subscriptions
send name and organization
to: [redacted] P14

P.L. 86-36

To submit articles or letters
by mail, to: P1, Cryptolog

via PLATFORM mail, send to:
cryptolg at barlcØ5
(bar-one-c-zero-five)
(note: no 'O' in 'log')

Editorial

This month we bid farewell to our assistant editor, [redacted] who is retiring (the verb, not the adjective).

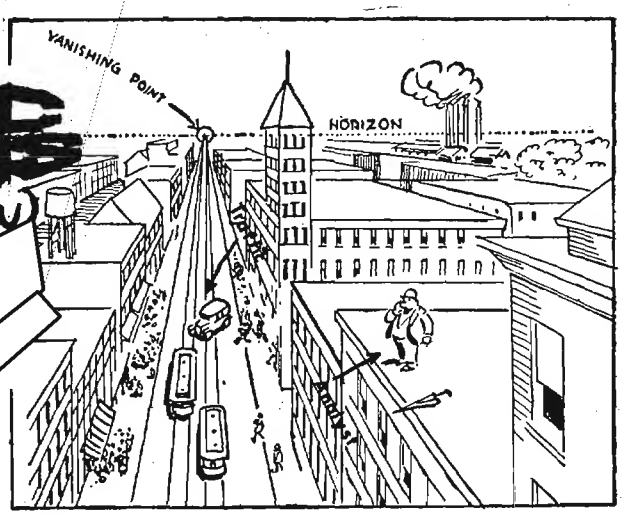
If you have ever known him, you will find it hard to imagine him as a retiree. Harry doesn't seem to have a "second gear" or any way to go at half speed. There is a rumor that when Harry talks REALLY fast, only dolphins can understand him.

We will miss him. He has been a stalwart force in keeping this magazine going.

Thanks, Harry, for your inexhaustible supply of energy, enthusiasm, and encyclopedic knowledge. Shalom.

TRAFFIC ANALYSIS: A CURRENT PERSPECTIVE (U)

by CW2 USAFS



~~(C)~~ Traffic analysis today is beset by a variety of problems affecting the analyst, not the least of which is a lack of understanding concerning the separate functions which comprise traffic analysis. Contrary to popular definition, the traffic analysis (TA) field of today encompasses much more than the mere analysis of communications externals. Traffic analysis consists of six separate functions:

[] Collection Management;

[] Collection Support;

[] Processing;

[] Analysis;

[] Reporting; and

[] Evaluation.

which are intended to be mutually supportive of each other. However, because both managers and analysts lack understanding of the interrelationship that exists between and among these functions, they all too often are performed at cross purposes, with one function inadvertently affecting the success (or failure) in a follow-on activity. In these days of meager collection and analytic resources, it is imperative that an understanding of the relationship bonding the six functions of traffic analysis be fostered at all levels to ensure the complete and proper use of these resources.

~~(C)~~ To function effectively in the traffic analysis world of today, the traffic analyst must have a rudimentary knowledge of several areas in addition to communications externals. Among these areas are

Collection Management: including the Collection Objective Performance Evaluation Systems (COPES) and its follow-on Collection Evaluation System (CES) statistical reporting;

Collection Support: including knowledge of working aids available to assist in collection and identification duties, as well as of target characteristics that can support both current and future collection operations;

Processing: including Automated Data Processing (ADP) routines available to assist in manipulating the intercepted data for subsequent analysis;

Analysis: including sufficient knowledge of the various traffic analysis techniques and target characteristics that develop information to satisfy consumer requirements;

Reporting: including the criteria and appropriate reporting vehicles for providing information to satisfy consumer requirements;

Evaluation: including the techniques available to constantly monitor and compare tasked and acquired collection so as to redirect these resources to areas of the target's communications that are susceptible to exploitation; and

Perspective: including a thorough knowledge of the target's order-of-battle (OB) structure that underlies the target's communications, as well as a thorough understanding of the targets past history from which to view changes in the target's communication habits.

A lack of knowledge in any one of these areas can ensure failure in attaining the ultimate goal of consumer satisfaction. Additionally, an incorrect managerial emphasis on any one function over another will also ensure failure in attaining this ultimate goal.

Collection Management

(U) The first function to be encountered by the traffic analyst is the Collection Management function. It is at this point that the consumer's requirement is translated into specific collection tasking to acquire the needed information.

Collection Support

(U) Once the tasking is accurately developed and transmitted to the field site, the next function, Collection Support, determines the success or failure of a field element in satisfying its tasked requirements.

(U) At the local level, field site managers usually are faced with a decision of where to use their most talented personnel: an effective Collection Support element or an effective Analysis & Reporting effort. Unfortunately, all too often the choice is for the A&R effort because of the high visibility of analytic breakthroughs and consistent product reporting.

(U) This decision on the part of the field manager leaves the Collection Support effort

with poorly trained analysts, generally those with the least experience in TA duties. These individuals are then placed in the fast-paced environment of the collection floor, an environment that breeds confusion and is the least likely place to learn the functions of traffic analysis.

(U) In addition to the assignment of poorly trained personnel at the local level, at the national headquarters level technical support material is seldom provided in a manner that facilitates its use in the field. Most often this material is seldom provided in a bulk manner, which dictates tedious manipulation in the field to extract desired information. Yet this same manipulation is often performed within the national headquarters on a daily basis through various computer routines but is not made available to field sites. Thus, while a steady and voluminous stream of technical support material is provided to the field element, it is seldom provided in a manner that facilitates its use in the field, and it is often too unwieldy to be used on the collection floor, with all its hectic activity.

Of these, the misidentification rate should be of the greatest concern since it dramatically impacts on the Processing, Analysis, Reporting, and Evaluation functions that follow the Collection Support function. A poor case (notation) identification rate in the Collection Support function means that all case identifications received for processing must be verified to ensure that the correct notation is applied, a time-consuming process for an individual already strapped for time in performing his or her other processing duties. Failure to verify the correct identification on all cases can have disastrous effects:

[] on processing, necessitating complete revisions of case history records;

[] on analysis, which attributes incorrect technical operating characteristics to the wrong case;

[] on reporting, which provides false information to the consumer; and

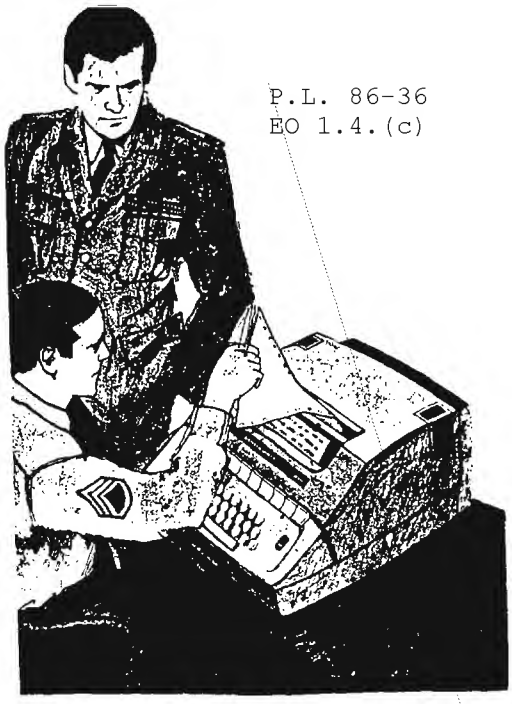
[] on evaluation, which attributes tasking satisfaction to the wrong target.

To ensure the most effective Collection Support effort possible, field managers must develop an operational system that exposes all elements of the assigned analytical work force to both collection support duties and desk analysis duties. Ideally, such a system, which rotates the individuals between these two worlds, will ultimately increase the expertise levels available in both areas. Likewise, at the national level all methods of developing collection support material for field elements must be closely scrutinized to ensure that a minimum of manipulation is required onsite to use the data.

Processing

~~(C)~~ The next function of traffic analysis, Processing, often assumes a life of its own. From the analyst's standpoint, processing (or, to use a more common term, "logging") quickly becomes associated with drudgery. The fact that processing is accomplished merely in order to organize intercepted data for subsequent analysis is quickly forgotten. Processing then becomes a mindless transfer of data from one medium, the raw intercept, to another medium, the casebook or data base. Further, this attitude quickly leads to overlooked items of significance that otherwise would require immediate analytic or reporting attention.

P.L. 86-36
EO 1.4.(c)



P.L. 86-36
EO 1.4.(c)

(U) From the managers' standpoint, however, processing quickly becomes the key productivity measurement because of the statistical data available from data base maintenance inputs. Since the underlying desire is for an accurate data base from which to extract material for subsequent analysis, and from which to drive any automatic collection support vehicles which emanate from the data base, the managers' concern becomes one of concern over data base maintenance accuracy and volume. Unfortunately, this concern for accuracy normally translates into a concern over the correct formatting of the input rather than over the validity of the information itself. Likewise, the volume concern translates into a meaningless "body count" that is used to spur further productivity.

[redacted]

poses should be deleted from existing requirements. For example, the processing of voluminous message serialization data should not be imposed on the affected analyst if no one intends to perform a study of this aspect of the target's communications. The processing requirements should be continually geared to the voids of the target's communications that need to be exploited, and not to those characteristics that are commonplace and seldom change, as these voids are reduced through exploitation.

P.L. 86-36
EO 1.4.(c)

Analysis

[redacted]

underlying military force must be the first priority of the traffic analyst. Only through such mastery can an analyst understand the communications facing him or detect changes in those communications which can be attributed to something reportable. In addition to ensuring that the analysts are knowledgeable of the structure of the military force, managers at all levels must ensure that the analytic studies to be undertaken first have the potential of satisfying either established reporting criteria or of developing material that will aid Collection Support activities. All other studies, no matter how well-intentioned, are merely a waste of an analyst's valuable time.

Reporting

(U) The reporting function is the sum total of how well the earlier functions were conducted. If these earlier functions were conducted in a shoddy manner, they normally will culminate in an inaccurate report that bears little relationship to the consumer's initial state requirement. Conversely, if these earlier functions were conducted in an orderly manner so that each function supported the following one, the report will probably satisfy the consumer's request. In addition to satisfying the consumer's request, the analyst must also concern himself or herself about using the proper format for the information to be used in the report. Poorly formatted reports, while factually correct, convey a sense of disorganization that can color the credibility placed on the information by the recipient.

(U) Both of these rationalizations have effectively downgraded the opportunities available to the traffic analyst to participate in the reporting function, the only function which at present is directly keyed to reporting criteria that match the consumers' requirements. By reducing the analysts' participation in this function, we also reduce their ability to ensure that the previous functions remain directed towards the ultimate goal of consumer satisfaction. Managers at all levels must emphasize the development and maintenance of effective and imaginative reporting programs that challenge the analysts' skill to ensure that that reporting function does not become divorced from all that has preceded it.

P.L. 86-36
EO 1.4.(c)

Evaluation

P.L. 86-36
EO 1.4.(c)

(U) Through their involvement in the evaluation process, the traffic analysts are well placed to redirect the collection resources to exploitable areas--or at least to areas that have not been satisfied through past intercept--of the target's communications. Unfortunately, the evaluation process is normally either not performed at all or it keys on indicators which, when viewed alone, fail to provide an accurate reflection of the productivity of the resource.

[redacted]

(U) As should be apparent, the traffic analyst of today is faced with a variety of functions. However, all of these functions are tied together by a thread of continuity relating to the final goal of consumer satisfaction. If these functions are not kept in their proper relationship to each other, either through analyst neglect or improper managerial emphasis, it will be difficult--if not impossible--to achieve a cohesive operational mission that satisfies consumer requirements. Only through mutually supporting functions, as originally envisioned, can our scarce collection and analytic resources be used to their fullest extent. In other words, perhaps a return to the basics may be in order for the traffic analyst of today!

P.L. 86-36
EO 1.4.(c)

To: Editor, Cryptolog

Dear Ed:

(U) "To err is human, to forgive--divine." We should feel that way, but it's not easy. When the Data Standards Center was alerted to look for that verbum horribile "JULIAN DATE" in the latest Shell Game article entitled "Time Shells" (February-March 1984, pp. 9-11) our dismay was great. Here again was that atrocious misnomer staring at us from the pages of an otherwise excellent publication. Only arch conservatives who don't believe in progress would dare use the term "Julian Date" when what they meant was really Ordinal Date (which is a Federal standard as well as DoD/NSA.) After all, when Pope Gregory XIII gave us the modern calendar (circa 1580), he likely never dreamed that 20th century man would still be using the JULIAN calendar after all these years. It's now 13 days behind our Gregorian one, which is hard enough to keep up to date, what with leap years and all that.

(U) So, please, repent and be forgiven--and henceforth use only ORDINAL DATE when you want to record a date such as 21 June 1985 in the form "85172." (The Standards Center expounded on this theme in the September '83 issue of CRYPTOLOG, in a piece called "Do You Really Mean Julian?") Verb. sap.

[redacted]
P13D

P.L. 86-36

Dear Editor:

Improved communications, processing centralization, and widespread computer use have all contributed to the apparent demise of the ITC (Informal Technical Circular) vehicle used by the traffic analyst. No one seems to remember when the last ITC was issued or who is the keeper of ITC serialization (or even if they are serialized for that matter). This is all very regretable because after many long years of searching I have finally found a TA item that meets ITC criteria and merits widespread circulation among the agency's analytic work force. As a last resort this ITC is being sent to you in the hope that it can be published via CRYPTOLOG.

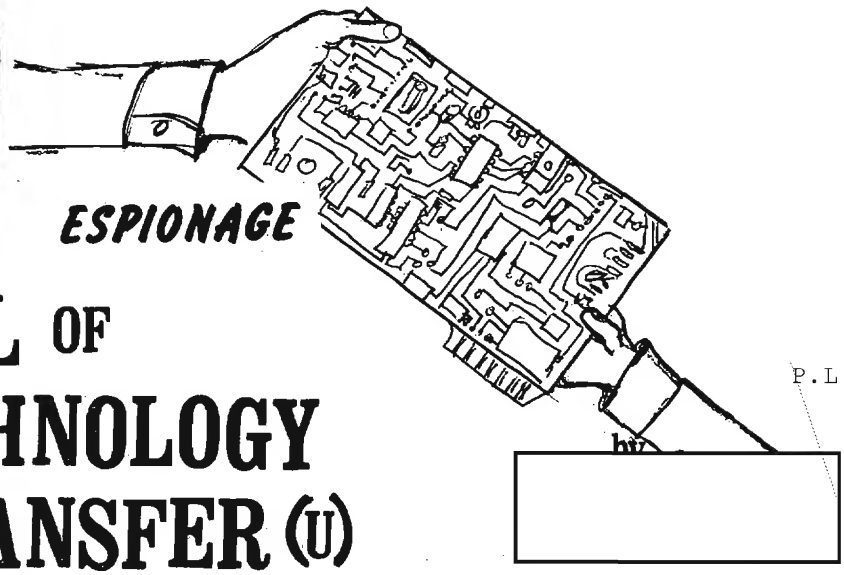
[redacted]

[redacted]

jim@barlc05

P.L. 86-36
EO 1.4.(c)

ESPIONAGE AS A TOOL OF TECHNOLOGY TRANSFER (U)



P.L. 86-36

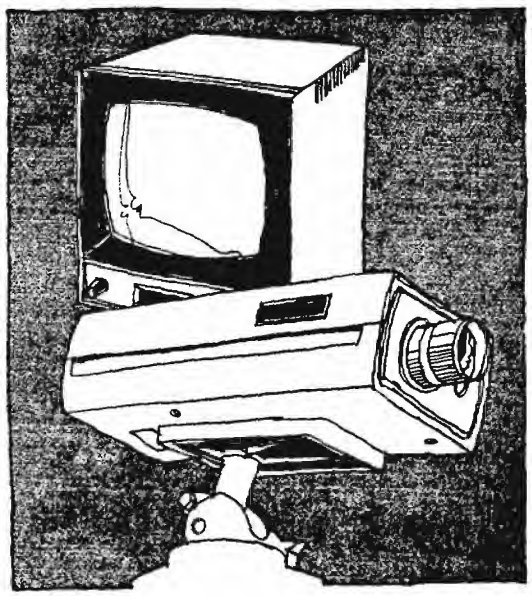
Technology has become one of the major variables that influences and determines United States national security policy. What can or cannot be done as a part of this policy is largely determined by the technological capabilities of the United States and of its adversaries. Technology thus has a major impact on strategy with tactical and strategic doctrine constantly changing in response to rapid advances in technology. These changes necessitate an almost continual alteration of strategy and policy to incorporate improvements in technology and to counter those of our potential adversaries. Our enemies are faced with the same dilemma.

The result of technology becoming a vital component of a nation's security is that any power that lags significantly in military technology, no matter how large its military budget or how efficiently it allocates resources, is likely to be at the mercy of its more progressive enemies. Thus the importance of military science and technology to national security is increasing in almost direct proportion to the importance of science and technology to society. Although the United States has enjoyed military superiority since World War II because of the advanced scientific and technical base that was built up over a number of years, today this dominance is being threatened by the Soviet Union.

The Soviet Union views scientific and technological achievements as the keystone of political and economic power; however, they traditionally have done well only in processes not noted for technological complexity in which central coordination of large amounts of capital and resources confer some kind of comparative advantage. This has caused them problems since there is more to technological innovation than the mere discovery of the concept. The concept has to be translated into something useful. Nonetheless, the gap between US and Soviet military technology is closing and we are rapidly approaching an era of equality. Although the Soviet Union is currently the second most technologically advanced nation in the world, we continue to provide them with the products and processes that they require to develop new and increasingly more sophisticated weapons systems.

Modern weapons become obsolete in a comparatively few years since the lead time to develop weapons is often considered long when compared to the rate of change in science and technology. The cost and complexity also increases with each new generation of weapons and this has resulted in a premium being added to careful and accurate research and development planning. Sudden technological achievements that impact on doctrine and set nations off on a new course are rare and occur only after decades of research and development. The development of these weapons systems is time-consuming, expensive, and important. Since the success of a country's foreign policy rests increasingly on technological superiority, technology has become one of the most decisive influences in international affairs. The necessity for obtaining technological information has become paramount to existence as a world power; therefore, espionage to determine where on the scale of technological possibilities the enemy lies has become important, commonplace, and necessary for the survival of many governmental systems.

The necessity for obtaining military and political information by espionage has always been accepted, although it has been the convention never to acknowledge that it was being done. As early as World War II the Soviets began to use espionage to improve their technology. The friendly relationships between the Allies during the war years made espionage easy and lucrative. As a result, the Soviets gained important knowledge of the US atomic energy program, data on jet engine propulsion, and access to many US production techniques. In addition, because of the Lend-Lease program, they had actual US equipment to use as models. Even so, despite a theoretical heritage and German scientists captured at the close of World War II, the Soviets began from a rather meager technical base. Until the early 1950s they continued to rely heavily on outside assistance. With the development of thermonuclear weapons and the systems to deliver them during the mid-1950s, scientific and technological advances became more rapid and more important. As a result, the Soviets increased and intensified their intelligence collection, using espionage to gain information regarding advances in science and technology. Acquiring or stealing technology was assigned the highest priority for Soviet intelligence operations. Both the KGB and GRU, the two major Soviet intelligence organizations, were tasked with using whatever means were available to ensure that a continuing flow of Western technology found its way to Soviet agriculture, Soviet industry, and, most important, to Soviet military users. Thus, the collection of scientific and technical information became a normal function of the Soviet intelligence services.



In the United States, more than three-fourths of all military-related research and development is done by private industry with the Department of Defense contributing over $200 million a year to their basic research activities. In addition, there are many other research activities that often result in new processes or techniques that could have military application. The difficulty is that a piece of research can serve both the purposes of peace and war. This causes considerable problems in controlling this kind of information since the openness of our society has caused a dilemma. In the name of national security, our government often feels compelled to constrain the flow of technology to the Soviet Union or other countries who might pass on the critical information to the Soviets or their counterparts. For economic health, however, the US government must promote the export competitiveness of our high-tech products. These controls have always been at the heart of controversy and have made the denial of information to the Soviets difficult. This has made the acquisition of information by their intelligence apparatus highly successful and therefore very damaging to the interests of the United States.

The research and development programming system of the United States and its annual military posture statement give many details of our future military plans. These documents are widely distributed and available to the Soviet agent. As stated in a 1970s study on the arms race, "The United States, for reasons of policy, tradition, and ineptness in keeping secrets, has been so open that Soviet military planners perhaps know as much about these matters in the US as the US military planners." Using this type of data, Soviet agents can be directed towards available open source documents. They find them in great numbers. It is estimated that there are at least 50,000 technical journals, most of them monthly, published throughout the world and that, in some technical areas, notably electronics, plastics, and chemicals, literature output doubles every 10 to 15 years. About 30 percent of the agents' requirements could be met by legal, open means such as subscribing to Aviation Week, attending international conferences, using contacts with Soviet citizens attending American universities, and through the adroit use of the Freedom of Information Act. Often by asking the right questions they are able to acquire from the federal government files and other technical data materials not generally available to the public. Much of this data is often recently declassified.

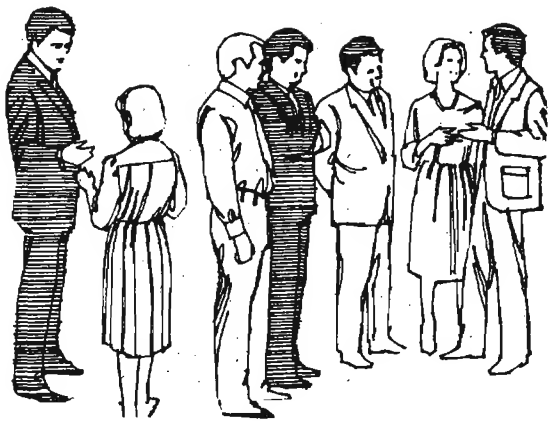
The Soviets are also able to buy technologically sophisticated equipments on the open

market. Legal acquisitions generally have their greatest impact on the industrial base and affect military technology on a relatively long-term basis. These legal acquisitions almost always find their way into military industries and subsequently into the civilian sectors of industry that support military production. Although the Soviets will state that these acquisitions will be used solely for civilian applications, legal purchases are part of a well-organized and well-coordinated plan to acquire Western technology that is militarily significant and of benefit to Soviet industries engaged in research, development, and production of weapons systems.

Although massive amounts of information are gained by using open source publications and through the legal acquisition of American equipment, about 70 percent of the Soviets' technology requirements are met by their illegal purchases and through information gained by their intelligence services. According to Edward J. O'Malley, an FBI official in Washington, "They buy what they can, they steal what they can't buy." [1] Clandestine acquisition of the West's most advanced military-related equipment and know-how by the Soviet intelligence apparatus is a major and growing problem. These organizations have been so successful that the manpower allocated to this effort has increased significantly and there are now several thousand technology collection officers operating in the United States under various covers ranging from diplomat to journalist to student. FBI Director William Webster, in an appearance on CBS's "Face the Nation" on 24 April 1983, stated that "there are about 3,000 Soviet bloc diplomats in the United States and 30 to 40 percent pursue US secrets--especially military information and laser and computer technology." A favorite technique of these agents is to set up a dummy corporation chartered to provide an authentic-looking letterhead. At the present time, the Soviets are bankrolling about 400 of these companies in the free world. These companies place orders for the high-tech equipment required by the Soviets. The equipment is then sent to some non-Communist country in Europe and eventually diverted to Russia. [2] Another method used by the Soviet agents is to persuade legitimate foreign companies to import US equipment and then sell it to the Russians for a profit. The acquisitions gained through these illegal trade channels often have important military applications.

If the Soviet agent fails to acquire the necessary equipment or information using the methods described above, he may resort to larceny. One method used successfully is to persuade a disaffected employee of a high-tech company, usually through monetary inducements, to sell the company secrets. If necessary, the agent may even hire thieves to steal the information. This practice has been particularly lucrative for thieves in the Silicon Valley, where many have found that stealing microchips is more profitable and less risky than stealing cars. Among some of the secrets recently lost to the Soviets through the illicit transfer of technology have been the details of the quite radar system for the B-1 and Stealth bombers, the F-15 look-down, shoot-down radar system, and the Phoenix air-to-air missile--clearly some of our most modern and most important military technology. So confident are the Soviets that they can obtain whatever US technology they need that they now start building the new weapons system before the necessary high-tech components have emerged from US laboratories--and promptly steal them when we do.

When the Soviets receive Western technology, whether it was obtained through legal or illegal means, it is coordinated with information and equipment obtained through a complex network of international governmental, scientific, and technical exchange agreements the Soviets maintain with industrial nations worldwide. These include know-how, equipment, and computer data base collection activities of Soviet scientists and engineers who participate in academic, commercial, and official science and technology exchanges. Richard D. DeLauer, Under Secretary of Defense for Research and Engineering says that one problem with these exchanges is that US scientists and engineers "go off in the corner and shoot the bull with the Russians and anyone else who's there for the wrong reasons." [3] These visiting Soviet technical and student delegations generally consist of expert scientists, many of whom are connected with classified work. The mastery they gain in assimilating one technology enables them to use subsequent transfers of related technologies. The increased mastery that results can be used to undertake independent technological efforts that may include replication and adaptation of foreign technologies and eventually the creation of new technology.



Soviet candidates in various academic and scientific exchange programs nearly always propose research activities that involve technologies that have a military application. This is especially true in areas where the Soviets have a deficiency. In the past two years, more than 30 percent of the proposals offered under the graduate student exchange programs have been unacceptable because of the prospective technology loss. [4] Many other proposals had to be modified before they could be allowed. These Soviet students are often intelligence agents with the necessary scientific and technological backgrounds to make them extremely efficient in gathering information on technological interest.

Soviets also regularly attend high-technology trade shows and attempt to visit commercial firms, particularly small and medium size firms that are active in developing new technologies. These activities are used as a subterfuge to gain access to emerging Western technologies before they have been identified by the US government as having military applications and to make contacts with the personnel of these companies for future exploitation.

The information-gathering efforts that have been described above save the Soviets billions in research and development costs. Stephen Bryen, Deputy Assistant Secretary of Defense for International Economic Trade and Security Policy says, "We keep getting numbers that are so big, and so frightening, that we have given up trying to calculate the amount. But it is in the billions." [5]

In a cost-benefit analysis, espionage may be the most economically productive element of the Soviet economy. The benefits, however, do not stop there and are not totally economic. Using US technology, the Soviets can develop countermeasures to negate our weapons superiority before the weapons are ever deployed. Since the US must rely on technological superiority to offset the Soviets' advantages in quantitative military power, the illegal transfer of advanced technology is providing the Soviets with the rope to hang us--exactly as Lenin prophesied over a half century ago. In addition, the Soviets learn from our mistakes. They can then focus their research and development capital to areas where we are the weakest and select from the best of both technological worlds. The Soviet use of industrial espionage imposes increasing costs on our economy and results in a continual struggle to overcome technology we have invented. As someone once said, "We have met the enemy and it is us."

Footnotes

1. "US Mounts a Belated Effort to Halt the Theft of Electronic Secrets," Life, April, 1983, pp. 29-36.

2. "FBI Claims 30 Percent of Soviets Here to Spy," Washington Times, April 15, 1972, p. 4.

3. "Padlocking the Laboratory," Business Week, April 2, 1983, p. 100.

4. Ibid.

5. "Stemming Flow of High Tech to East," Christian Science Monitor, 11 April, 1971, p. 4.

SOLUTION TO NSA-Crostic No. 55

"The Things They Say," by Doris Miller, KEYWORD, August 1968, reprinted in CRYPTOLOG, November 1976.

"'I know it doesn't make sense, but that's what it says!' This is the granddaddy of them all, the great classic disclaimer, ...the most spontaneous, universal, and irrepressible outcry in the translating world; surely there is no one in the business who hasn't given tongue to it at some time."

INTELLIGENCE ANALYSIS IN THE COMPUTER AGE (U)

by Jack Gurin, SRL

In intelligence work, the day of the 3 x 5 card file is over, or nearly so. Those files, on which so much depends, have moved into something more distant from the user--somehow more difficult to reach--an entity that sometimes seems to thumb its nose at those who would seek its help, the data base. But no matter what your opinion may be of the data base, whether you describe it with kind words or harsh ones, that is where you must go for the information you need. And what is it but a collection of all sorts of data used by many people for many different purposes. It should never be an end in itself, since its only reason for existing is to serve the needs of those who feed it and ask it questions. But, like some of the ancient gods, it sometimes assumes a role in which ritual supplants substance, and the needs of those who support it are subordinated to the requirements of the system created to serve them.

(U) The data base is a creature of the computer. Until recently, it has been possible to obtain information from this repository only by means of special languages and, even then, in a highly restricted number of approaches. As the computer becomes more versatile and responsive, it is not asking too much to expect the data base to move under the direct control of the user of the information, without requiring the acquisition of yet another strange language. With the rapid changes in computer technology and with the increasing accessibility of personal computers, it is time for us to take another look at data bases and how we may make them serve our purposes more effectively than in the past.

FILE SYSTEMS AND DATA BASE SYSTEMS

(U) Files have always been with us and probably always will be. There certainly will be a role for them in a highly localized arena. When the same type of data is needed in many different places by many different people, however, it becomes difficult to ensure that all copies of the information are accurate and up to date. Especially in a rapidly changing intelligence environment, the importance of a maintaining the accuracy, timeliness, and completeness of a commonly held fund of information is obvious.

(U) Another shortcoming of file systems is their inflexibility. If you need the data items grouped in different ways, there may be difficulties, delays, or denials. How often we have been disappointed because the organization of the data did not permit a particular question to be answered? And restructuring the data may be out of the question, since a seemingly trivial change in a file environment sets off a chain reaction of other changes that must be made. This is the kind of thing that gives data systems a bad name, since they seem not to be able to fulfill their purpose without requiring expensive alterations every time a slightly different question is asked.

(U) A file is a thing unto itself. It just isn't tied to anything else, any more than a book is tied into another book. Of course you may have cross references to other files, just as you may find a find a bibliography in a book you are reading. If the cross references or the bibliographies are extensive, they discourage the user from any attempt at com-

pleteness. Unless you have lots of time, it is not practical to follow all the suggested additional references.

THE HIGH COST OF COMPUTER PROGRAMMING

(U) It is all to obvious that programming costs have consumed a larger and larger part of the data processing budget and, as far as the user is concerned, most (or all) of that has been spent on keeping the old systems going, not improving them but keeping them alive. Often it is the user who triggers many of these costs by asking for information that is just a bit different than what can be derived from the system as it stands. This kind of proliferation of maintenance requirements, if not carefully controlled, tends to increase as the number of programs grows, until it threatens the entire programming budget by its size.

(U) One of the main objects of a data base system is to allow changes in the information to be made just once, and to have that update available immediately to all users. Another is to permit a programmer to make a change in one of the elements of a data base without having to change the other structures. In other words, the data should be independent of the programs as far as the user is concerned. To achieve these goals we have to move into data base management systems that have been created for just these purposes. It is not necessary for the user to know how a data base management system works; it is enough to know that it exists only to provide, easily and quickly, the needed information (if it is stored within the system).

DATA BASE DIALOGUES

(U) The user should be able to deal with the data base without relying on intermediates. Data should be organized in such a manner that, by employing a simple, easy-to-use query language, the user will be able to express needs directly and extract satisfactory responses. The job of combining those elements in the data base that must be brought together to provide a complete and accurate answer to the query is a task for the data base system, not for a special program that needs to be written for the purpose.

(U) No doubt when this kind of flexibility is provided, much more use will be made of the data bases, especially as the users learn more about the potential of such a system in the exploitation of accumulated information.

SIZE OF THE DATA BASE

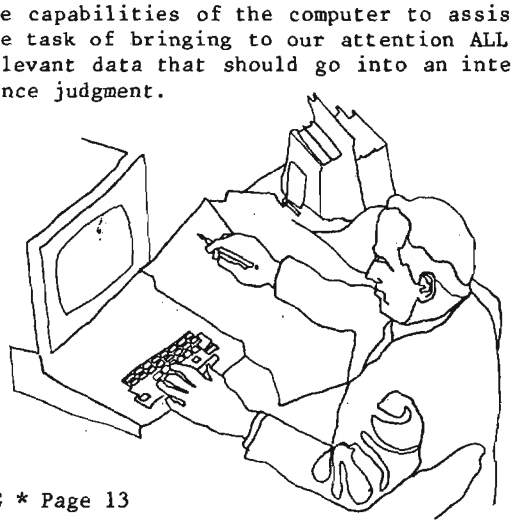
(U) When we talk of THE data base, it should not be assumed that there should be one humongous collection of information, with all users dipping into the reservoir when looking for answers. What appears more likely is that there will be many data bases, all on the same system, but physically and logically separate. They should be linked by common query languages and design policies so that, when necessary, linkages between them may be forged. It is entirely possible that the system will be called upon to provide information that transcends the province of any one collection of data, no matter how comprehensive it may be.

P.L. 86-36

DATA SHARING

(U) Ignoring for the moment the obvious considerations of security, which will vary from one application to the next, one of the great advantages of effective use of a data base system is sharing the data. This means that everyone who is in the system and needs it has access virtually immediately to the information in its most recent and accurate form. It will be necessary to develop clear lines of responsibility (where they do not already exist) as to which is the office of primary interest for each type of data item and to ensure that all participants understand the distribution of responsibilities for entering and maintaining the data items.

(U) The difficulties involved in linking together or integrating the data base and of standardizing the data items may turn out to be considerable, since so many organizations have developed their own ways of keeping files of information. But that is a cost that must be borne if we wish to take full advantage of the capabilities of the computer to assist in the task of bringing to our attention ALL the relevant data that should go into an intelligence judgment.



THE DATA MODEL

(U) Data items do not exist in isolation but are associated with one another. Maps need to be drawn showing which data items are associated with which others, and what types of association these are. Such maps give an overall representation of the data that is needed to service the inquiries of users. A map showing different data items and how they are associated with one another is called a data model. The development of such a data model, or map, is the first step in organizing a data base that will be most responsive to the user's requirements.

(U) Development of a model calls for the direct participation of users, since they are the ones to define the various elements that are to go into the data base, and they must also describe the nature of the relationships between those elements. Some kind of consensus must be reached if the model is to be suitable for a group of users, but in all cases the opinions of the user rather than the data system professional must prevail.

(U) The user's view of the data may be much simpler than the actual data and should be tailored to his own application. This structure may be referred to as the "logical" structure. The actual data structure, stored on tape or disks, may be referred to as the "physical" structure. To illustrate this difference, think of the map of the Metro system in Washington as it is displayed in the subway cars. It merely displays the relationships of the stations to each other and to the line they are on. There is little representation on that schematic of actual distances between stations and actual paths followed by the train. The map is a logical rather than a physical map of the system. The user of the system, whether subway or data base, need not be concerned about anything but the logical structure. When using the data base, nothing about the system should remind him that the physical structure is different. It will be up to the system to convert the user's queries, whether naive or informed, into whatever is required to provide the desired response.

CREATING THE DATA MODEL

(U) The question of what would be the best logical structure for the data is vitally important for the success of the data base system. We expect the system not only to provide complete, rapid, and accurate responses to today's queries from the user, but also to allow itself to be updated easily and, as intelligence needs vary in the future, to be responsive to changes in the nature of queries. If the logical structures are designed badly, operational personnel will not rely on them but will turn once again to local files which, for all their shortcomings, can be created and manipulated with some feeling of security.

(U) Another key factor in determining the success of the data base is the extent to which the user can be protected from the impact of changes that have to be made in the structure. These changes can be minimized by careful system design, but they are unlikely to be eliminated. When such changes are to be made, they should be accomplished in a manner that does not alter the external appearance of the system. The data may be expected to change constantly--especially in an intelligence environment--but the structure of the data base itself should be as stable as possible. If in the future our data base system itself will require the services of programming personnel to make repeated changes in order to cope with changing demands, we will return to the same old problem of expending an unconscionable proportion of resources in patching up the system to keep it going. Chances are that only the minimal investment will be made, and the system will creak badly.

(U) When talking about data modeling, we refer to items about which we store information as entities. Examples of entities are personalities, weapons systems, organizations, abbreviations, job titles, and abstract concepts.

(U) Each entity has various attributes that contribute to the definition of that particular entity. It may be a rank, a location, a title, or other type of data. Of course what is an attribute in one case could well be an entity in another. If one is pursuing a personal name, then organization becomes an important attribute. If, on the other hand, an organization is the focus of interest, then a membership list could be an important attribute. The same case could be made for the relationships between, say, pilot, aircraft, and flight number. The challenge to developers of a data base system is to make the most effective use of the entity-attribute relationship.

FOURTH GENERATION LANGUAGES

(U) As was said earlier, the data base should be so configured that the user will be able to deal with it directly, seldom having to rely on a programmer to adjust the system to his changing requirements. One of the obstacles facing many users or would-be users of computer-based information systems has been the mysterious language that must be mastered before one is allowed to talk to the computer.

Over the years, the number and levels of sophistication of such languages have increased, and we are at the point where the machine can be addressed in words that almost resemble (but really aren't) normal English. Some of these languages are procedural and some non-procedural; some have both capabilities. A procedural language specifies how something is accomplished; a non-procedural language specifies what is to be done but not in detail how it is to be done. For example, should you wish to go by taxi to BWI Airport, you would expect to enter the cab and simply announce your destination to the driver, confident that he would know how to get there. You made a non-procedural request. If, however, he did not know the route, you might have to describe not only which roads to take, but when to turn, etc. In the latter case, you would specify the how as well as the what and would issue procedural instructions. To illustrate this point with a query directed to the computer, suppose your question was "Give me the names of all pilots stationed at Andrews AFB who are qualified to fly F-16s." With a procedural language you would have to specify the path to follow:

```
get personnel list for officers at Andrews AFB
find rated officer
     select if qualified to pilot F-16
     do until no more rated officers
print names of selected officers
end
```

When all the conditions for our ideal data base are fulfilled, the user will merely say what is to be done and will not be concerned with how. For the latter he will rely on the application generator, a part of the system that is designed to perform that task.

(U) Users should not only be allowed to create their own applications without having to resort to programmers, but they should also be able to direct the system to extract wanted data and to format it into reports. Some of these report generators are independent of the data base or query facilities; others are extensions of the data base query languages. Ideally the user should be able to start by learning to make simple data base queries and should steadily extend that skill to data manipulation and report formatting.

(U) Another facility available to some data base users permits them to specify WHETHER they want the data displayed in graphic form. Then they specify HOW. The system may be asked to search files or data bases and to chart the resulting information according to different criteria.

(U) All of the preceding marvels are not only possible. They are reality in some lucky circles. Before they can be made available locally, however, the system must be redesigned and equipped with data base management facilities suited for creating applications. A better logical design is required for the data bases now in use. Only then can effective use be made of data base query languages, report generators, and application generators. Remember that we are talking not about what may be possible in the future, but about the present state of the art. And it is likely that the coming months will bring additional advances that will permit the user to manipulate data even more effectively and speedily.

A NEW ERA FOR INTELLIGENCE ANALYSIS

(U) What is the significance of these advances in data base manipulation to translators, transcribers, analysts--anyone engaged in intelligence analysis? Why go through the trouble--and it will require quite a bit of effort--to reorganize all information files according to a demanding discipline that will appear foreign to previous practice and experience? What kind of payoff is to be expected from this investment of time, labor, and sophisticated new equipment?

(U) The justification lies in the potential of the computer to serve the intelligence community as it has the scientific and technological communities. The computer will be able to assist the user of information files with greatly increased speed, accuracy, and completeness, and this alone could raise productivity sufficiently to justify the investment. In addition, the computer will be able to assist the user by exploring the relationships of data items in a way that the analyst could only do if time and effort expenditure limitations did not exist. Data base searches of possible relationships, performed at enormous speed, could provide to the user information that would otherwise simply not be available. The combination of a powerful personal computer and a relational data base that may be reached via a non-procedural language and a data base management system will provide a tool for intelligence analysis such as we have never had before.

NAMING SOVIET CITIES (U)

Since the October Revolution, the Russians have seen an inordinate number of their larger and more important cities undergo name changes at least once, and, in some cases, as many as three times. A study of this twentieth-century phenomenon reveals certain trends which have developed in Soviet city name-changing through the years. In addition to isolating such trends, a study of this type gives us a quick course in Soviet history over the past fifty years or so. In the material which follows only the larger Soviet cities have been included.

(U) Not long after the Bolsheviks assumed power a crusade was started to change the names of Russian cities which were named after tsars or tsarinas, or were otherwise associated with tsardom. Accordingly, Novonikolaevsk became Novosibirsk, Aleksandrovsk was changed to Zaporozh'e, Ekaterinoslav was renamed Dnepropetrovsk, Tsarkoe Selo became Pushkin, etc. Today none of the major Soviet cities have Tsarist-associated names although there is still a fair smattering of Ekaterinovkas, Nikolaevkas, Elizavetovkas, etc., throughout the country.

(U) Another noticeable trend in city naming began with the death of Lenin: important Soviet cities were renamed for deceased Bolshevik heroes. It was perfectly fitting that after Lenin's death, his birthplace and the place where the Bolshevik Revolution erupted should be named for him. Other leading Bolsheviks like Sverdlov, Kirov, Kujbyshev, and Frunze were also posthumously honored by having cities named after them.

(U) Reprinted from the Spring 1984 issue of Vox Topics, where it appeared with no author's name given. City names were written with Cyrillic characters in the original and the only change made in the text was to transliterate those names for readers of Cryptolog.

(U) During 1924, the very year of Lenin's death, another less durable city-naming practice, that of naming large cities after living Soviet heroes, came into vogue. Cities all over the Soviet Union began taking Stalin's name. Zinoviev also had this dubious honor bestowed on him when Elizavetgrad became Zinovievsk. Later Molotov and Voroshilov were similarly honored when the cities of Perm' and Lugansk respectively were named after them. The trouble with the practice was that most of these men fell into disfavor while still alive or, in the case of Stalin, after death. This necessitated changing the names of the cities once again. Rather than dreaming up new names, the Soviets tended to resolve these dilemmas by giving the cities back their harmless former names. Thus, Perm' and Lugansk where reinstated while Dyushambe, which had become Stalinabad, returned as Dushanbe. The only two major cities to escape this fate were Kalinin, formerly Tver', and Gorkij, formerly Nizhni Novgorod [1].

(U) New names, however, had to be found for Zinovievsk and Stalingrad which could not go back to their tsarist-tainted named of

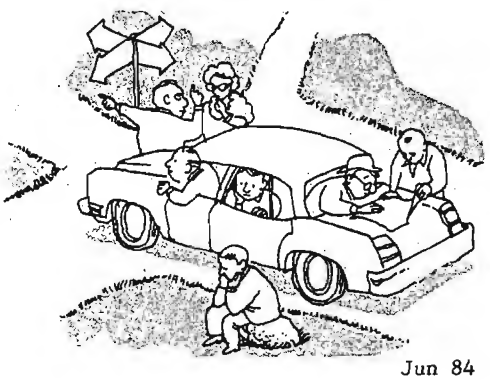
Elizavetgrad and Tsaritsyn. The same was also true for Stalino which could not return to Yuzovka, which had been named after an Englishman. In searching for new names for these last two cities, the Soviets decided to name them after fairly noncontroversial things--rivers. Thus, Volgograd and Donetsk came into being. Zinovievsk had its name changed twice more, first to Kirovo and finally to Kirovgrad.

(U) The Soviets in their naming and renaming of cities over the years seem to have become more conservative and sophisticated. No major cities have been named after living persons in recent times. It seems apparent that they have finally realized the problems related to bestowing names of living Soviet giants on their major cities. They have carried this even further by not naming any larger cities after dead stalwarts in recent times. When the cosmonauts Gagarin and Komarov died there was no rush to name Kiev, Rostov, or other large cities after them. An interesting sidelight is that in the last few years the very few non-major Soviet cities to be named after dead people were named in honor of recently deceased foreign Communist leaders. Thus, Stavropol in Kujbyshevskaya Oblast' was named after the Italian Communist Togliatti, Chistyakovo took the name of the Frenchman Thorez, and Liski became Georgiu-Dej after the Romanian leader.

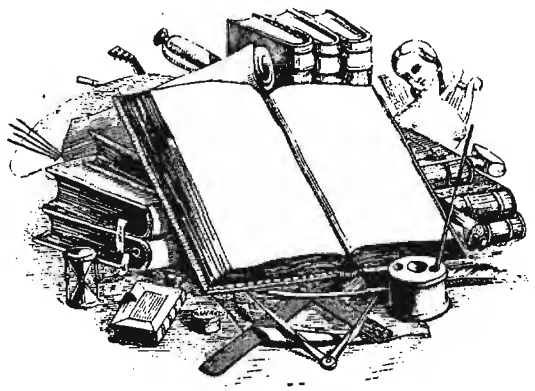
(U) The trend today, however, seems to be to hold on to the city names they have and to honor their heroes by naming schools, ships, parks, and the like after them.

---

1. Kalinin was a noncontroversial symbol of the Russian peasantry who got along with everyone, while Gorkij's literary fame evidently overshadowed any political troubles in which he became involved.



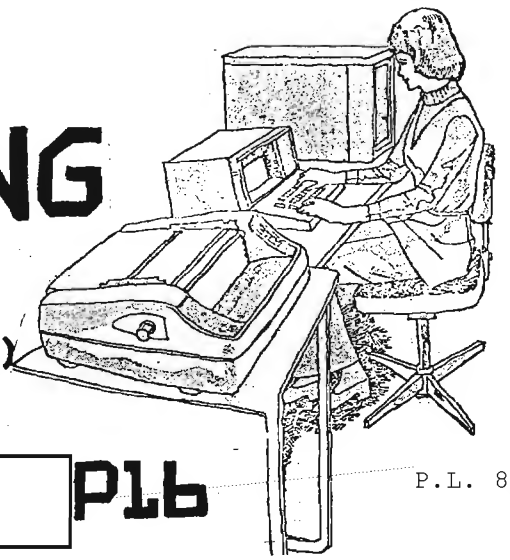
BOOKBREAKER (U)



excerpt from
Bookbreaking Improvisations, Part II,
by Stuart Buck
(P1 Informal, June 1966, p. 98)

WORD PROCESSING PLAIN & SIMPLE

(For the User, That Is) (U)



by [redacted] P1b

P.L. 86-36

(U) By now, most people have heard of word processors, and a fairly large segment of the Agency population has used one. We would here like to examine their characteristics and describe our "ideal" word processor. If in some small way our words help influence the features available to us, we shall be most grateful.

(U) The term "word processing" is a gift from the IBM Corporation coined, according to the Encyclopedia of Computer Science and Engineering,[1] in 1964, and like many products of that ubiquitous company it has--as they say--"caught on." Caught on so well, in fact, that it hardly seems necessary today to define it. Define it we shall, however, being ever solicitous of our readers and wanting to have some way to start. As we see it, a word processor is a program, system, or machine

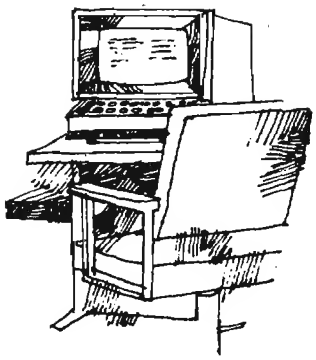
- [] which offers the capability to place a text onto an electronic medium, usually using a keyboard;
- [] which displays the text selectively to the user; and
- [] which offers a variety of features which facilitate modifications of the text.

Word processing is the process of doing all that.

(U) What, then, is the difference between word processing and text editing? One might well wonder. We have attempted mightily to find a consistent distinction, to no avail. "Text editing" is an older term, dating from a time when "terminal" referred to a printing device that might have a keyboard, as opposed to today's view that it is a scope. Therein lies much of the difference between the two usages, but there are pockets of vendors and users who make other, less consistent, distinctions between the two.

(U) The reason for the great popularity of word processors is that they take a lot of the pain out of writing--make it almost effortless, in fact. They are helpful in getting words recorded in the first place, and they come into their own when it comes to modifying the text or performing certain simple checks on it. They also have their little aggravations, not least of which, in our view, is the fact that none of them seems to have all the features we would like to have available to us.

(U) Popular as word processors have become, however, the novice user may have to overcome certain apprehensions before approaching an electronic device for the purpose of writing a few words. That tiny screen, for example, can't possibly hold the entire text (in fact, sometimes no more than half a typed page, and usually no more than two typed pages). It is helpful to think of the screen as a window onto the text. The user is given some method for navigating the text--that is, moving the window about so it offers views of other parts of the text. Once mastered, the process is no harder than flipping the pages of a book. A somewhat new concept is that of coming to terms with the computer about the location for some action to occur. If we wish to write on paper, we have only to position our pen appropriately for the writing to occur where we wish. For most word processors, we are not writing directly on the screen, so the computer must know where we want to write, and we must have confirmation that the computer knows that is where we want to write. This little problem is usually solved by the identification of a special symbol (underline, reverse blank, arrow, etc.) which marks our spot. This symbol is called a cursor.



Standard Features

(U) There are certain features that are common to most word processors, and we shall begin by reciting them as a means of getting them out of the way. We are pleased to see that the number of features a word processor is expected to support is growing, truly reflecting our desires.

(U) Most word processors can be directed to insert automatically sequenced page numbers, prepare headers and footers automatically (either may include the page number), justify either margin or both, and center text on a line. The cursor's movement about the screen is typically controlled by depressing a key (e.g., carriage return, tab, right-arrow) or moving a little device called a "mouse" (it's about the size of the animal and a long "tail" connects it to the computer). The user may intersperse the text with typed formatting commands, or by typing a function key, or by selecting a task from a "menu" of possible tasks, perhaps accompanied by manipulation of the mouse. Some word processors perform the command immediately, giving the user a "what-you-see-is-what-you-get" view of the text. Others embed the commands and require that the user perform a subsequent processing step in order to achieve the proper format. Moving the viewing window about in the text may be performed by means of function keys that cause the portion of text shown to go forward or back a designated amount; to the very beginning or very end, or to a specified line; or shift the window to the right or left. Alternatively, the user may move a graphical representation of a thumbnail to a location judged to be near the desired text, perhaps moving a cursor (the cursor causing text to be "scrolled", or moved line-at-a-time, up or down on the screen when it reaches the top or bottom, respectively, of the window) to reach the precise text segment. A desired location may also be reached by searching for a specific string of letters.

(U) More than one portion of the same or different texts may be placed on the screen simultaneously by means of multiple windows, also called "ports," that may be placed on the screen as desired. The overall capacity of the screen never changes, of course, so there is a logical limit to the number of windows that can be used realistically.

(U) Now we get to the part of the word processor that makes it all worth while. Once some text has been placed in the computer, all sorts of modifications can be made. A text sequence can be copied from one document to another, thus making it possible to create boilerplate files which may be copied selectively and repeatedly for the creation of a new document tailored to the needs at hand--a feature particularly useful for legal documents and the like. Instead of copying the text, the user may delete it from one part of the document and move it to another--a cut-and-paste type of procedure particularly useful for draft revisions. Or a string of text may be deleted completely. Word processors with a "spell" feature may be used to check each word of text against the system's dictionary, a very useful feature.

(U) Most word processors have a string search feature, which allows the user to command the computer to locate a selected string of characters. We might, for example, ask the computer to find the next occurrence of the string "thu" in this text. It would examine each of these characters in turn, until it reaches the end of the text or finds a match. If no match is found, the system so informs the user: if one is found, it moves the cursor to mark the string, thus.

Other Features

(U) Competition and success have produced the desirable effect that new and handy features are introduced regularly. The undesirable consequence is that the present-day user usually has to wait for another implementation of the system at hand before being able to enjoy the benefits. Some of the more recent features are designed to make data entry, as distinct from text modification, faster and easier.

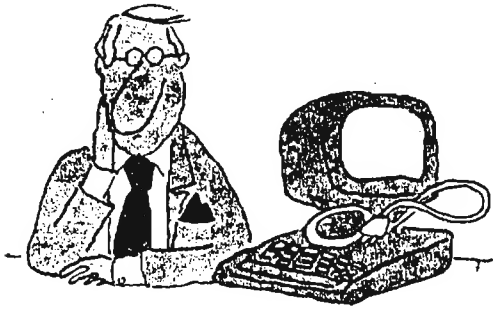
(U) "Word wrap" is a feature that has by now become almost universal. Word wrap occurs when the cursor reaches the right end of the screen. Instead of requiring the user to hit a carriage return at the appropriate place, the word processor positions the cursor at the beginning of the next line, pulling the unfinished beginning of the word along with it. For example, suppose the user has just typed "unfi" when the end of line is reached. When an "n" is typed, the word processor deletes "unfi" from the current line, places "unfin" on the next line, and positions the cursor immediately following the "n".

(U) Because margin justification is so commonly available, a good syllabification scheme is sorely needed. We have seen manuscripts that might have looked very professional because of the justified right margin but instead looked amateurish because there were too many spaces between words. This happens when there are a lot of long words in the text and the word processor has no capability to divide a word between two lines. If the final product is to look professional, syllabification, properly done, is de rigueur.

(U) The ability to change location of the margins, such as for a block indent is certainly a wanted feature.

(U) Some word processors give the user the option of marking certain words to be selected for the automatic preparation of an index. Some provide for special-purpose formats, such as an outline format, so that the user need only keep track of the level of indentation, and the system keeps the numbering scheme in sync. Others, such as those intended for use by computer programmers, identify structural errors in the text. Facility with manipulating columnar text is another feature of interest. Selection of special fonts, such as bold face and italics, or even foreign character sets, is another option. A few word processors offer a graphics capability.

(U) Selection of the font (e.g., bold face, italic, Times Gothic, Greek) and size of print are very desirable features for professional-looking copy. An appropriately adaptable printer is, of course, required to take advantage of such features.



(U) Many desirable features are needed frequently by some users but never by others. Among these are

[] superscripts or subscripts;

[] a decimal tab, which aligns a column of figures at the decimal point; and

[] an overstrike capability, which allows an accent mark to be placed above a letter or a sequence of text to be marked for deletion (useful especially for legal texts where it is important to display proposed textual deletions).

Some users will need a mail merge capability for sending out personalized form letters.

(U) We have not exhausted the list of useful features, but we have grown weary of the task and fear that an exhaustive list may, in any event, be virtually impossible to construct.

Some Cautions

(U) An apparent deficiency of a spelling checker cannot be attributed to the word processor but to its user. Until word processors with ESP are developed, there is no way for them to assure that what the user intended to say is represented accurately and completely. A spelling checker helps the user find words that are misspelled. Some spelling checkers perform better than others, but it is still the responsibility of the user to proofread the copy. At the very least, a sentence parser would be required to identify an "and" where you wanted an "an" or to spot that familiar problem of duplicated words, as in

Paris
in the
the Spring

or ungrammatical but properly spelled constructs. The phrase "Man bite dog" is not grammatical, but all the words are spelled correctly. The proper procedure, then, is to obtain a list of misspelled words, use the search feature to find the offenders, make the necessary corrections, and then review the entire corrected text. A powerful way of correcting the same mistake every time it occurs is to use the global-search-and-replace feature. Note once again, however, that it is not foolproof: the user who habitually transposes the same letters, rendering, for example, "the" as "hte", may want to change all misspellings throughout the text--without also creating a misspell in words like "heighten."

Problems

(U) There are obvious faults with a two-pass process, which requires the user to look at a file containing formatting instructions, then checking the formatted file to be sure everything is as it should be. Perhaps it is less obvious that the "what-you-see-is-what-you-get" approach may, unless very skillfully designed, introduce its own variety of difficulties. Suppose, for instance, we have a long text with many indented paragraphs interspersed with text at a standard line length. When entering this for the first time we want a quick and easy way to move back and forth between the various formats. Having to reset left margin, then reset right margin, and maybe also reset tab stops is not quick and easy enough for us. A "ruler line" that operates quickly and easily, and that the program knows not to print out on our copy, may do the job if easily accessed. Suppose that when we start to print this text, we discover that the line lengths are not suitable after all, and we must change them. We would like to be able to issue a global command that changes all the line lengths appropriately throughout the document (no protests, please: we are confident that with sufficient ingenuity, someone can satisfy our demands).

(U) Established typing and printing conventions are difficult for word processors, partly because their designers are unacquainted with the rules, and partly because completely correct implementation is difficult. Consider, for instance, the standard convention that two spaces separate sentences. When the word processor "fills" a line (brings text from the next line up to fill empty space in the current line), it typically separates the two pieces of text with a single space. If the last character is a period, however, it presumably must insert an extra space. Not all periods signal the end of a sentence, however. In the by-line of this article the reader will find a period that does not mark the end of a sentence. It is usually the case that a period preceded by a capital letter does not finish a sentence. An exception is when the author begins discussing, for example, a PhD. A question mark also identifies the end of a sentence and must be followed by two spaces. Complicating the issue are conventions regarding parentheses and quotation marks. A printing convention is to place a period within the final quotation mark, even when it is a "sentence-ender." Therefore, the word processor would need to check on the character preceding a quotation mark to determine whether or not it is a period and take appropriate action. (Parentheses, too, may enclose a sentence-ending period.)

(U) Other conventions which are frequently ignored by word processors are prohibitions against widows and orphans. A widow is a short line ending a paragraph which is carried over to the top of the next page or column. An orphan is the opening line of a paragraph dangling by itself at the bottom of a page or column. The general rule taught in typing classes is to have no fewer than two lines of a paragraph on a page. Similarly, no word should be divided between lines if the division leaves only two characters on either line. Hyphenation of the last word on the page is also considered bad form.

(U) Justification of both margins calls into focus certain subtleties of the printer's art. The importance of syllabification for an aesthetic appearance was mentioned earlier. Syllabification rules in English are quite complex, however, so designing a system which correctly chooses "prof-it" but "pro-fane," "la-tent" but "lat-i-tude," and "na-tion" but "nat-u-ral" is a nontrivial task. Scanning for "rivers" is another task of the printer that your usual word processor has not yet addressed to our satisfaction. As the reader surveys a page of text, there should be no sequences of white space (caused by inserting extra blank space between words) running across several lines and causing the reader to be distracted. Like housework, attention to small details like this is never noticed until neglected. It should be noted that until word processors made it easy to justify both margins, typeset materials had a corner on the market, and they are usually set with variable-width fonts. The advent of large amounts of monospaced text may accustom us to

its face, but we are not at all certain that we will ever like it as much as "properly" printed text.

(U) Some systems approach the notion of word wrap without giving the full-blown feature. Some versions of the Rand Editor, for example, position the cursor on the next line without dragging the unfinished word down with it. This defeats the intention of word wrap, since it saves neither the key stroke to reposition the cursor (carriage return) nor the effort of having to look at the screen to assure that the word is not split between lines.

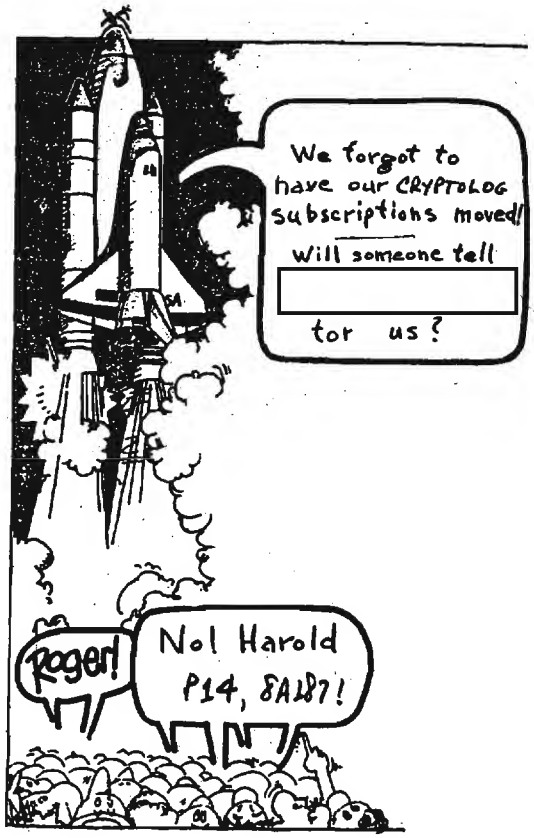
(U) Capital letters pose a lot of problems for word processor designers. For example, the search function is usually designed to find only an exact match, not allowing for a change in case. But the spell program may ignore capitalization and return to the user a list of not-found words in lower case only. Suppose it failed to find "capital" in its dictionary and we wanted to find it in our text. If we give that string to the search routine of most word processors, we will not find the first word of this paragraph.

(U) We have encountered other problems with the search function as well. Suppose the spelling checker produces in its list of unfound words the sequence "ns". Gazing at this sequence, we fail to remember anything we wrote that might have been represented thus, and we haven't a clue where in our 50,000-byte text this string may appear. We wish to find it and a long list of such strings as expeditiously as possible so we can print the text within the next hour. Since "ns" is a common sequence we prefer not to have to look at every instance of it in our text. So we try looking for " ns "; we have reasoned that if the word processor identified it as a word it will be bounded by spaces. Unfortunately, this is not necessarily the case. For most word processors once in each line the space between words is replaced by a carriage return. Therefore, if our "ns" falls at the beginning or end of a line, the match will not occur and the search fails. We may then try " Ns ", " ns", " Ns", "ns ", "Ns ", and "NS", perhaps not in that order. (If the spelling checker distinguishes between upper and lower case, we are spared some of these trials.) By now we have probably become quite dispirited and may no longer care whether we find the error or not. If we subsequently remember that we referred to "n's", which the spelling checker contracted to "ns" (now we remember), our humor is not restored.

The Ideal Word Processor

(U) We view our specification of the ideal word processor as an ongoing project. As soon as a word processor meets all our specifications, we are quite prepared to think up new ones. The system designer should not be discouraged by this prospect but rather should look upon it as job security if not as a challenge. In any event, let the record show that our Ideal Word Processor is not intended as a be-all-and-end-all, forever after, Amen. It merely reflects our thinking on the matter at this moment. We anticipate, that once you readers have been stimulated by these few notions, you will contribute many additional features which we will instantly laud and add to our list. Consider this, then, the beginning of our specification statement.

(U) As you might expect, we want all the desirable features now available to us. Put them all in the same package please. We tire of being teased by having word wrap but not font selection in one word processor while another one has font selection but not word wrap.



P.L. 86-36

(U) We of course want corrections to all the problems noted above. For the print conventions, we feel these have been stated adequately already, so we won't repeat the requirements. The search feature should be enhanced in several ways. We would like the option of finding a string exactly as given or with upper/lower case equivalents of letters (e.g., if the string is "next", having it find "Next" as well). When doing a search-and-replace in such instances, the usual need would be to have the replacement character be in the same case as the one replaced (e.g., replace "naxt" with "next" would replace "Naxt" with "Next"). We also wish to include a space in our search and retrieve any character which may serve as a word divider. This includes the carriage return character. Speaking of which, we may as well request the option to search for a function character.

(U) Cursor movement is another good candidate for enhancement. If a word processor is supposed to process words, why not let the user move the cursor through the text in word increments? We would like to move the cursor to the beginning of the next word, sentence, or paragraph--and, of course, also to the previous word, sentence, and paragraph. Another feature that would help the user who hits the wrong key or simply has a change of mind would be to send the cursor back to its location before the previous cursor change command. We cannot tell you how many times we have hit the "home" key by accident when we were busy placing the cursor just where we wanted it.

(U) There now, that wasn't so bad, was it? We would be ever so grateful if you could just see to it that we find this package all wrapped up in our next word processor.

NOTE

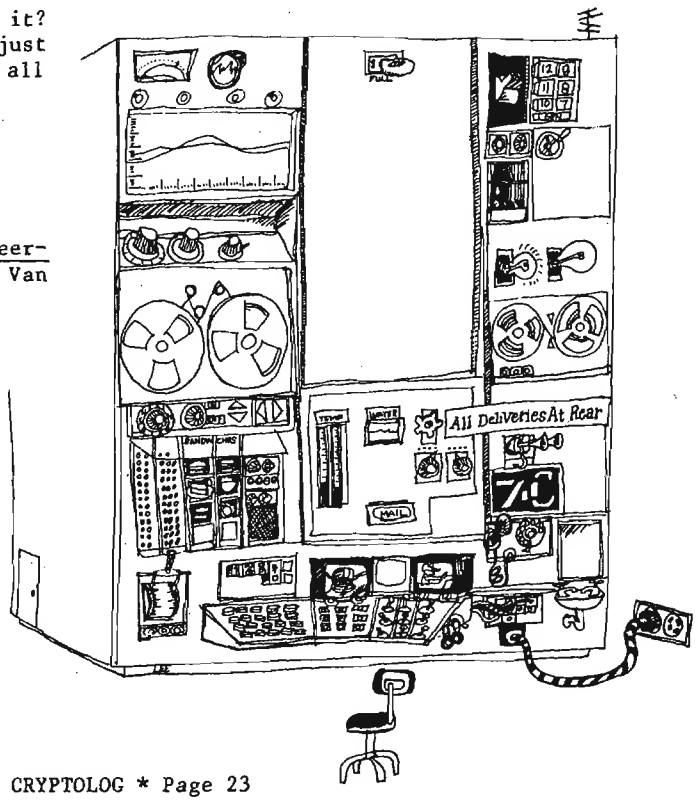
Encyclopedia of Computer Science and Engineering, 2d edition; Anthony Ralston, editor; Van Nostrand Reinhold Co., 1983.

Date: 6 Jul 1984 at 0943-EDT
From: vag at punix [redacted]
Subject: Editorial Reply P.L. 86-36
To: cryptolg at barlc05

Ekcej moernb asodfc dfk cmdf youdf wls c df basdfasdf?!?!?!?! Edf ckd sdfk tdsfyr df glcdoe? Thsas mrwo sdfkwe, cme pcm th wjch riddmre.

In other words, yes we do read the Editorial!!

[redacted]
T3511 P.L. 86-36



ANSWERS TO "WHAT'S THE CAPTION?"

From [redacted] P14: P.L. 86-36

"Operations Building 2A gargoyles near completion."

or

"Identify the seven staff officers in this picture."

or

"The tech track."

From [redacted] G31: P.L. 86-36

"I should have known there was a catch to his 'end-of-the-month special' on nose jobs."

From [redacted] B413: P.L. 86-36

Friar Nip: What the devil is Brother Chip up to?
Friar Tuck: His doctor told him to take two aspirins and gargoyle!

Friar Nip: That certainly is a marble-ous piece of work!
Friar Tuck: Yes, but future generations probably will take it for granite.

Friar Nip: What a farce! The sculptor has little talent, the stone is cheap, and the model is ugly as sin.
Friar Tuck: Maybe that's why Brother Chip has named his work "The Statue of Limitations."

From Clint Brooks, S4:

"SRB Program Review"

From "Jess Possible"

"I can see where the badge goes, but where do we punch in the numbers?"

Here's an extract from a collection of papers that ought to be distinguished in some way from the usual Golden Oldies. It's a prediction that came true, though it seemed far-fetched until just a few years ago. V.V.

Perhaps the principal element in the ignorance of the public in cryptography has been the attitude of the military and diplomatic governmental authorities of the world. It has not been deemed wise to admit the public to a knowledge of a science which might at some time or other endanger the safety of a state. Governments have gone so far as to restrict the use of secret means of communication to its diplomatic or military agents. A ban of this kind was placed by several of the European states in the 18th and 19th centuries. With the advent of more democratic forms of government, the public has taken to inquiring more closely into the affairs which concern it. Peoples everywhere have begun to insist that they be admitted into the confidence of the governments which rule them. It is to be expected, therefore, that the science of cryptography which has hitherto been so jealously guarded by a chosen few of the elect will in a few more years become a matter of more general knowledge.

extract from "General Principles of Cipher"
by James Rives Childs,
1st Lt. Inf, U.S.R.
April 1919.

Unless Texts Hang Together, Linguists Will All Hang Separately or Coherence, Cohesion, & LG-140 (U)

by [redacted]

P.L. 86-36

(C) In his article in the December 1983 Cryptolog entitled "Non Posse vs. Posse Non (U)," [redacted] brings up a point that is fundamental to Agency translators and, in many cases, transcribers as well. Communications, be they in our personal lives or those of our target countries, have one very important fact in common: there is always a message. It so happens that in order for the message to be understood, there are certain conventions that must be followed by both sender and receiver. Our knowledge of these conventions or characteristics of messages allows us to process the many messages we receive every day.

(U) The problem is that with regard to our native language, we are not aware of our use of these message traits for our understanding because they were acquired along with the language that we learned as children. Moreover, the large majority of language teaching separates the language from its main purpose, that of sending messages. Put another way, the method of teaching isolated vocabulary items, word endings, sentence, patterns, etc., puts the emphasis on the medium rather than the message.

[redacted]

P.L. 86-36

[redacted]

(U) Granted, an experienced trainer can shorten the process somewhat, but what is often the case is that the experienced linguist has internalized those tools he uses to deal with a message. As a result, it is as hard for him to pass these tools on to the new linguists as it is for us to explain a certain quirk in English grammar to a non-native speaker. In light of this situation, how does the new linguist learn the point that Mr. [redacted] has "been trying to push over the years: all parts of a text are related!" without a long period of trial and error?

(C) The National Cryptologic School has one answer in the form of a new course, LG-140, Applied Cryptologic Linguistics. This class is focused directly on the problem of how we process language at the Agency. While it does cover some linguistic theory and some descriptions of various types of language applications at NSA, the majority of the course consists of lectures, exercises, and class discussions about the characteristics of messages or texts and how this nature can be exploited.

(U) The course covers such characteristics of texts as COHERENCE and COHESION, to name just two. Cohesion, the property of a text to hang together syntactically, is illustrated by various kinds of chaining exercises where one follows a certain element of a text from the beginning to the end, despite the fact that it takes various forms throughout. One chaining example focused on the different ways of

be certified by their parent U.S. ...ment organization to NSA Visitor Clearance (M5611) at lea... ...king days prior to the date of the visit.

(2) The NSA/CSS sponso... ...at M5611 is on distribution for all incoming clearance... ...tion messages to include SOCOMM Mess... ...at are recei... ...the sponsor. When M5611 ... is not indi... must provide a copy c... M5611 fo...

(3) The NSA... Request (Form G2450) for eac... submitted to Vi... Clearance (M... cation at leas... ...days prior ... M5611 maintai... ...on i... cleared and SI...

(4) Vis... concerning foreign nation... Directorate of Foreign Relations ...

(5) ... concerning congr... the Legislative ...

(6) Addi... controls, and sponsor's responsibilities is set ... Chapter 803, NSA/CSS PMM 30-2.

b. ~~(FOUO)~~ Visits to Field Activities and to Non... Facilities

(1) The certification of appropr... access information is required when NSA/CSS organizations or facilities on classified ... certifications will be limited to require ... ...ivities of the U.S. Government. The clearance/ac... ...ation is accomplished as follows:

(a) CONUS - NSA/CSS Form G2901, "Official Visit Notification", is prepared in the sponsoring NSA/CSS element and forwarded to Visitor Clearance (M5611). M5611 provides the necessary certification to the organization or activity being visited. Guidance for the preparation of Form G2901 is contained in Annex A to this Letter.

(b) Overseas - Planning messages, which are prepared within the sponsoring element for release by Travel Management and Support, M62, must contain appropriate security clearance/access authorization information.

referring to a Metrobus in an article about a bus accident. Although this may seem a trivial example, it nonetheless illustrates something that can be very problematic, especially in voice texts.

(U) Coherence, often lumped together with Cohesion, is really a measure of how well those concepts which are elicited by the text match our understanding of the world. Put another way, Coherence is a measure of the degree to which a text makes sense. Something can be cohesive (or syntactically correct) but still not make sense. One place this often occurs is in newspaper headlines, such as

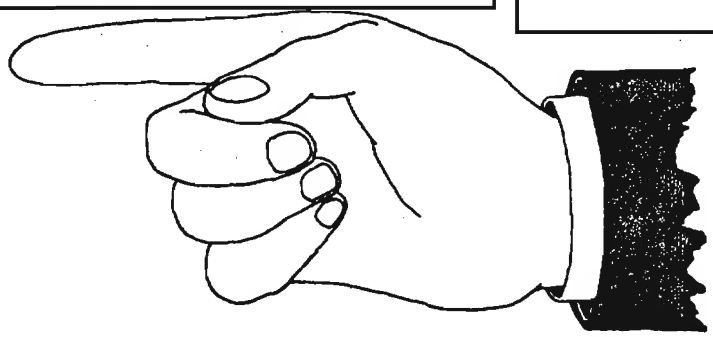
HERSHEY BARS PROTEST

EO 1.4.(c)
P.L. 86-36

or

ESCAPEE RECAPTURED IN SANDWICH

Both of these examples (which were taken from LG-140 material) are grammatically correct, but it is not until we apply our world knowledge that they really make sense. You have to know that General Lewis Hershey, the head of Selective Service, refused to let some draft protestors carry out their activities near his headquarters and that Sandwich is a town in Massachusetts.



NSA-Crostic No. 56

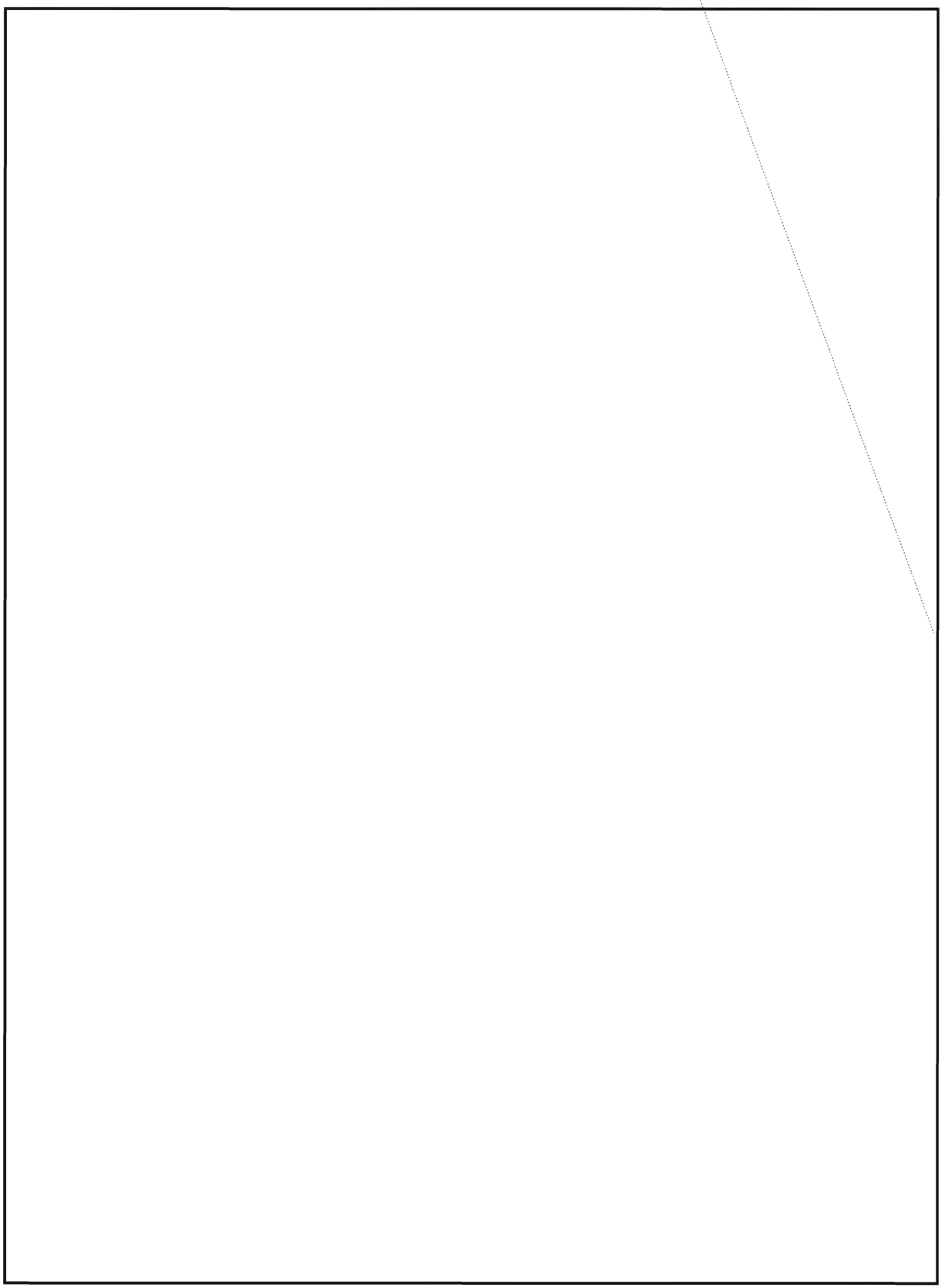
THE FEMALE LEADS OF WORDS H, J AND S

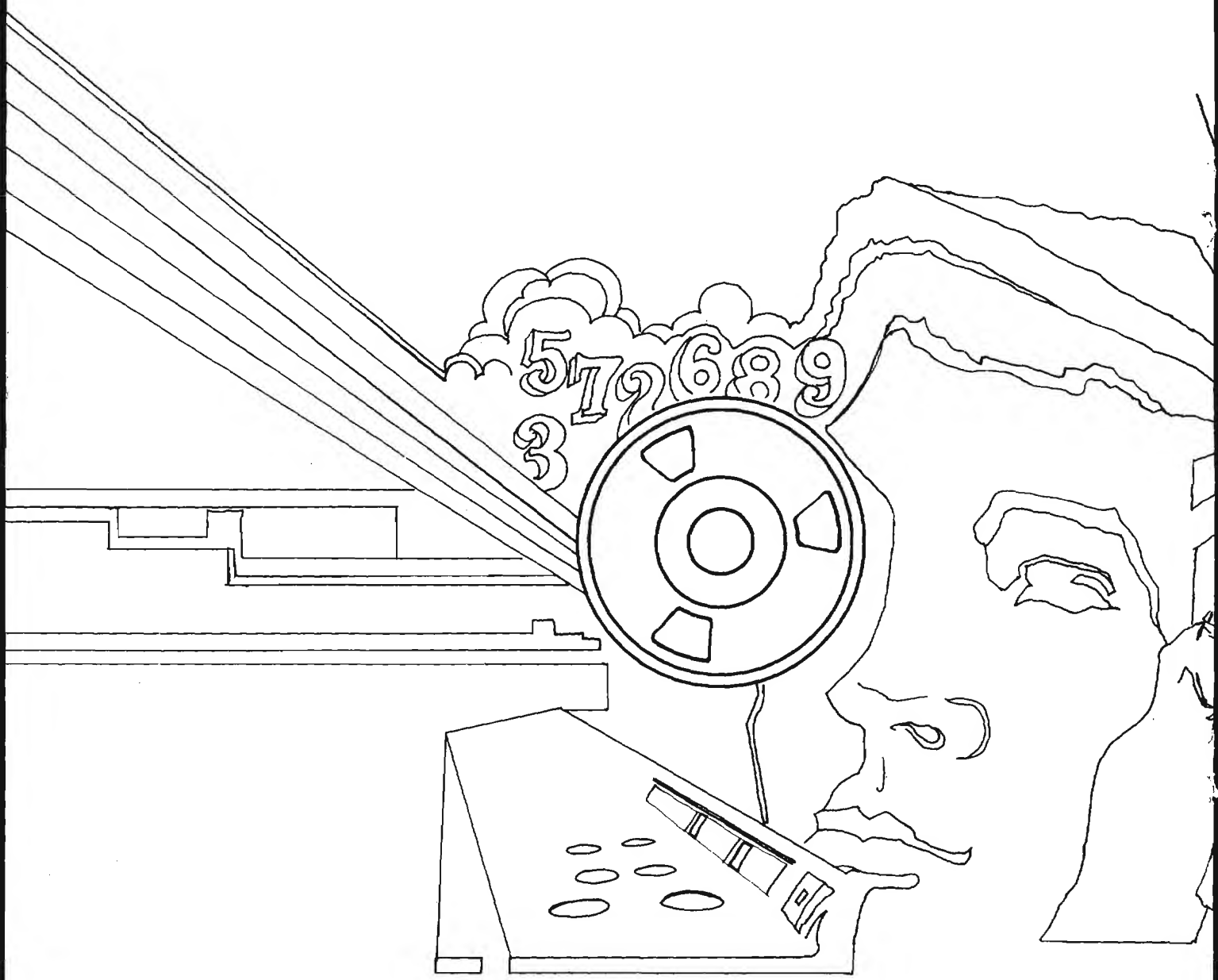
by [redacted]

P.L. 86-36

[redacted]

P.L. 86-36




5726893